

1600

CRF Errors Edited by the STIC Systems Branch

Serial Number: 10/017,736C

CRF Edit Date: 2/12/04
Edited by: AS

ENTERED

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

___ Corrected the SEQ ID NO. Sequence numbers edited were:

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

___ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

___ Moved responses to same line as heading/numeric identifier, specifically:

✓ Other: Sequence 1- inserted hard return after "1056"



1600

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:49

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

4 <110> APPLICANT: Thibeault, Diane
 5 Lamarre, Daniel
 6 Maurice, Roger
 7 Pilote, Louise
 8 Pause, Armin
 10 <120> TITLE OF INVENTION: Purified Active HCV NS2/3 Protease
 13 <130> FILE REFERENCE: 13/082
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/017,736C
 C--> 15 <141> CURRENT FILING DATE: 2001-12-14
 15 <150> PRIOR APPLICATION NUMBER: 60/256,031
 16 <151> PRIOR FILING DATE: 2000-12-15
 18 <160> NUMBER OF SEQ ID NOS: 25
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1230
 24 <212> TYPE: DNA
 25 <213> ORGANISM: HCV
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (1)...(1230)
 31 <400> SEQUENCE: 1
 32 atg gac cgg gag atg gct gca tgc tgc gga ggc gcg gtt ttc ata ggt 48
 33 Met Asp Arg Glu Met Ala Ala Ser Cys Gly Gly Ala Val Phe Ile Gly
 34 1 5 10 15
 36 ctt gca ctc ttg acc ttg tca cca tac tat aaa gtg ctc ctc gct agg 96
 37 Leu Ala Leu Leu Thr Leu Ser Pro Tyr Tyr Lys Val Leu Leu Ala Arg
 38 20 25 30
 40 ctc ata tgg tgg tta cag tat tta atc acc aga gtc gag gcg cac ttg 144
 41 Leu Ile Trp Trp Leu Gln Tyr Leu Ile Thr Arg Val Glu Ala His Leu
 42 35 40 45
 44 caa gtg tgg atc ccc cct ctc aat gtt cgg gga ggc cgc gat gcc atc 192
 45 Gln Val Trp Ile Pro Pro Leu Asn Val Arg Gly Gly Arg Asp Ala Ile
 46 50 55 60
 48 atc ctc ctc acg tgc gca gtc cac cca gag cta atc ttt gac atc acc 240
 49 Ile Leu Leu Thr Cys Ala Val His Pro Glu Leu Ile Phe Asp Ile Thr
 50 65 70 75 80
 52 aaa ctc ctg ctc gcc ata ttc ggt ccg ctc atg gtg ctc cag gca ggc 288
 53 Lys Leu Leu Leu Ala Ile Phe Gly Pro Leu Met Val Leu Gln Ala Gly
 54 85 90 95
 56 ata acc aaa gtg ccg tac ttc gtg cgt gcg cag ggg ctc att cgt gcg 336
 57 Ile Thr Lys Val Pro Tyr Phe Val Arg Ala Gln Gly Leu Ile Arg Ala
 58 100 105 110
 60 tgt atg ttg gtg cgg aag gct gcg ggg ggt cat tat gtc caa atg gcc 384

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:49

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

61 Cys Met Leu Val Arg Lys Ala Ala Gly Gly His Tyr Val Gln Met Ala
62      115      120      125
64 ttc atg aag cta gct gcg ctg aca ggt acg tac gtt tat gac cat ctc 432
65 Phe Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His Leu
66      130      135      140
68 act cca ttg cag gat tgg gcc cac gcg ggc cta cga gac ctt gca gtg 480
69 Thr Pro Leu Gln Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala Val
70 145      150      155      160
72 gcg gta gag ccc gtc atc ttc tct gac atg gag gtc aag atc atc acc 528
73 Ala Val Glu Pro Val Ile Phe Ser Asp Met Glu Val Lys Ile Ile Thr
74      165      170      175
76 tgg ggg gcg gac acc gcg gca tgc ggg gac atc att tca ggt ctg ccc 576
77 Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Ser Gly Leu Pro
78      180      185      190
80 gtc tcc gct cga agg gga agg gag ata ctc ctg gga ccg gcc gat aat 624
81 Val Ser Ala Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Asn
82      195      200      205
84 ttt gaa ggg cag ggg tgg cga ctc ctt gcg ccc atc acg gcc tac tcc 672
85 Phe Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ser
86      210      215      220
88 caa cag aca cgg ggc cta ctt ggt tgc atc atc acc agc ctc aca ggc 720
89 Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly
90 225      230      235      240
92 cgg gac aag aac cag gtc gag ggg gag gtt caa gtg gtc tcc acc gct 768
93 Arg Asp Lys Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr Ala
94      245      250      255
96 aca caa tct ttc ctg gcg acc tgc gtc aac ggc gtg tgt tgg act gtc 816
97 Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr Val
98      260      265      270
100 ttc cat ggc gcc ggc tca aag acc ttg gcc ggc ccc aaa ggc cca atc 864
101 Phe His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile
102      275      280      285
104 acc cag atg tac act aat gtg gac cag gac ctc gtc ggc tgg cag gcg 912
105 Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln Ala
106      290      295      300
108 ccc cct ggg gcg cgc tcc atg aca cca tgc acc tgc ggc agc tcg gac 960
109 Pro Pro Gly Ala Arg Ser Met Thr Pro Cys Thr Cys Gly Ser Ser Asp
110 305      310      315      320
113 ctc tat ttg gtc acg aga cat gcc gac gtc att ccg gtg cgc cgg cgg 1008
114 Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg
115      325      330      335
117 ggc gac agt agg ggg agc ctg ctc tcc ccc agg cct gtc tcc tac ttg 1056
118 Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu
119      340      345      350
121 aag ggc tct tcg ggt ggc cca ctg ctc tgc cct tcg ggg cac gct gtg 1104
122 Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala Val
123      355      360      365
125 ggc atc ttc cgg gct gct gtg tgc acc cgg ggg gtt gca aaa gcg gtg 1152
126 Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala Val

```

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:49

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

127      370      375      380
129 gac ttc ata cct gtt gag tct atg gaa act acc atg cgg act agt agc 1200
130 Asp Phe Ile Pro Val Glu Ser Met Glu Thr Thr Met Arg Thr Ser Ser
131 385      390      395      400
133 gct tgg cgt cac ccg cag ttc ggt ggt taa 1230
134 Ala Trp Arg His Pro Gln Phe Gly Gly *
135      405
138 <210> SEQ ID NO: 2
139 <211> LENGTH: 409
140 <212> TYPE: PRT
141 <213> ORGANISM: HCV
143 <400> SEQUENCE: 2
144 Met Asp Arg Glu Met Ala Ala Ser Cys Gly Gly Ala Val Phe Ile Gly
145 1      5      10      15
147 Leu Ala Leu Leu Thr Leu Ser Pro Tyr Tyr Lys Val Leu Leu Ala Arg
148      20      25      30
150 Leu Ile Trp Trp Leu Gln Tyr Leu Ile Thr Arg Val Glu Ala His Leu
151      35      40      45
153 Gln Val Trp Ile Pro Pro Leu Asn Val Arg Gly Gly Arg Asp Ala Ile
154      50      55      60
156 Ile Leu Leu Thr Cys Ala Val His Pro Glu Leu Ile Phe Asp Ile Thr
157 65      70      75      80
159 Lys Leu Leu Leu Ala Ile Phe Gly Pro Leu Met Val Leu Gln Ala Gly
160      85      90      95
162 Ile Thr Lys Val Pro Tyr Phe Val Arg Ala Gln Gly Leu Ile Arg Ala
163      100     105     110
165 Cys Met Leu Val Arg Lys Ala Ala Gly Gly His Tyr Val Gln Met Ala
166      115     120     125
168 Phe Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His Leu
169      130     135     140
171 Thr Pro Leu Gln Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala Val
172 145     150     155     160
174 Ala Val Glu Pro Val Ile Phe Ser Asp Met Glu Val Lys Ile Ile Thr
175      165     170     175
177 Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Ser Gly Leu Pro
178      180     185     190
180 Val Ser Ala Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Asn
181      195     200     205
183 Phe Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ser
184      210     215     220
186 Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly
187 225     230     235     240
189 Arg Asp Lys Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr Ala
190      245     250     255
192 Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr Val
193      260     265     270
195 Phe His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile
196      275     280     285
198 Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln Ala

```

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:49

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

199      290      295      300
201 Pro Pro Gly Ala Arg Ser Met Thr Pro Cys Thr Cys Gly Ser Ser Asp
202 305      310      315      320
204 Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg
205      325      330      335
207 Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr Leu
208      340      345      350
210 Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala Val
211      355      360      365
213 Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala Val
214      370      375      380
216 Asp Phe Ile Pro Val Glu Ser Met Glu Thr Thr Met Arg Thr Ser Ser
217 385      390      395      400
219 Ala Trp Arg His Pro Gln Phe Gly Gly
220      405
223 <210> SEQ ID NO: 3
224 <211> LENGTH: 1011
225 <212> TYPE: DNA
226 <213> ORGANISM: HCV
228 <220> FEATURE:
229 <221> NAME/KEY: CDS
230 <222> LOCATION: (1)...(1005)
232 <400> SEQUENCE: 3
233 atg aaa aag aaa aag ctc gag cat cac cat cac cat cac act agt gca 48
234 Met Lys Lys Lys Lys Leu Glu His His His His His Thr Ser Ala
235 1 5 10 15
237 ggc ata acc aaa gtg ccg tac ttc gtg cgt gcg cag ggg ctc att cgt 96
238 Gly Ile Thr Lys Val Pro Tyr Phe Val Arg Ala Gln Gly Leu Ile Arg
239 20 25 30
241 gcg tgt atg ttg gtg cgg aag gct gcg ggg ggt cat tat gtc caa atg 144
242 Ala Cys Met Leu Val Arg Lys Ala Ala Gly Gly His Tyr Val Gln Met
243 35 40 45
245 gcc ttc atg aag cta gct gcg ctg aca ggt acg tac gtt tat gac cat 192
246 Ala Phe Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His
247 50 55 60
249 ctc act cca ttg cag gat tgg gcc cac gcg ggc cta cga gac ctt gca 240
250 Leu Thr Pro Leu Gln Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala
251 65 70 75 80
253 gtg gcg gta gag ccc gtc atc ttc tct gac atg gag gtc aag atc atc 288
254 Val Ala Val Glu Pro Val Ile Phe Ser Asp Met Glu Val Lys Ile Ile
255 85 90 95
257 acc tgg ggg gcg gac acc gcg gca tgc ggg gac atc att tca ggt ctg 336
258 Thr Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Ser Gly Leu
259 100 105 110
261 ccc gtc tcc gct cga agg gga agg gag ata ctc ctg gga ccg gcc gat 384
262 Pro Val Ser Ala Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp
263 115 120 125
265 aat ttt gaa ggg cag ggg tgg cga ctc ctt gcg ccc atc acg gcc tac 432
266 Asn Phe Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr

```

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:50

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

267      130      135      140
269 tcc caa cag aca cgg ggc cta ctt ggt tgc atc atc acc agc ctc aca 480
270 Ser Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr
271 145      150      155      160
273 ggc cgg gac aag aac cag gtc gag ggg gag gtt caa gtg gtc tcc acc 528
274 Gly Arg Asp Lys Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr
275      165      170      175
277 gct aca caa tct ttc ctg gcg acc tgc gtc aac ggc gtg tgt tgg act 576
278 Ala Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr
279      180      185      190
281 gtc ttc cat ggc gcc ggc tca aag acc ttg gcc ggc ccc aaa ggc cca 624
282 Val Phe His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro
283      195      200      205
285 atc acc cag atg tac act aat gtg gac cag gac ctc gtc ggc tgg cag 672
286 Ile Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln
287      210      215      220
289 gcg ccc cct ggg gcg cgc tcc atg aca cca tgc acc tgc ggc agc tcg 720
290 Ala Pro Pro Gly Ala Arg Ser Met Thr Pro Cys Thr Cys Gly Ser Ser
291 225      230      235      240
293 gac ctc tat ttg gtc acg aga cat gcc gac gtc att ccg gtg cgc cgg 768
294 Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg
295      245      250      255
297 cgg ggc gac agt agg ggg agc ctg ctc tcc ccc agg cct gtc tcc tac 816
298 Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr
299      260      265      270
301 ttg aag ggc tct tcg ggt ggc cca ctg ctc tgc cct tcg ggg cac gct 864
302 Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala
303      275      280      285
305 gtg ggc atc ttc cgg gct gct gtg tgc acc cgg ggg gtt gca aaa gcg 912
306 Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala
307      290      295      300
309 gtg gac ttc ata cct gtt gag tct atg gaa act acc atg cgg act agt 960
310 Val Asp Phe Ile Pro Val Glu Ser Met Glu Thr Thr Met Arg Thr Ser
311 305      310      315      320
313 agc gct tgg cgt cac ccg cag ttc ggt ggt aaa aag aaa aag taa 1005
314 Ser Ala Trp Arg His Pro Gln Phe Gly Gly Lys Lys Lys Lys *
315      325      330
317 ggatcc 1011
320 <210> SEQ ID NO: 4
321 <211> LENGTH: 334
322 <212> TYPE: PRT
323 <213> ORGANISM: HCV
325 <400> SEQUENCE: 4
326 Met Lys Lys Lys Lys Leu Glu His His His His His His Thr Ser Ala
327 1      5      10      15
329 Gly Ile Thr Lys Val Pro Tyr Phe Val Arg Ala Gln Gly Leu Ile Arg
330      20      25      30
332 Ala Cys Met Leu Val Arg Lys Ala Ala Gly Gly His Tyr Val Gln Met
333      35      40      45

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/017,736C

DATE: 02/12/2004
TIME: 17:18:51

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\02122004\J017736C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:19; Xaa Pos. 6

Seq#:20; Xaa Pos. 6

VERIFICATION SUMMARY

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 17:18:51

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\02122004\J017736C.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1077 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1080 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1084 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1085 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:1103 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:20
L:1104 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0



1600

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 11:19:37

Input Set : A:\13_082 Substitute Sequence listing US.txt

Output Set: N:\CRF4\02122004\J017736C.raw

4 <110> APPLICANT: Thibeault, Diane
 5 Lamarre, Daniel
 6 Maurice, Roger
 7 Pilote, Louise
 8 Pause, Armin
 10 <120> TITLE OF INVENTION: Purified Active HCV NS2/3 Protease
 13 <130> FILE REFERENCE: 13/082
 C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/017,736C
 C--> 15 <141> CURRENT FILING DATE: 2001-12-14
 15 <150> PRIOR APPLICATION NUMBER: 60/256,031
 16 <151> PRIOR FILING DATE: 2000-12-15
 18 <160> NUMBER OF SEQ ID NOS: 25
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0

ERRORED SEQUENCES

22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1230
 24 <212> TYPE: DNA
 25 <213> ORGANISM: HCV
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (1)...(1230)
 31 <400> SEQUENCE: 1
 32 atg gac cgg gag atg gct gca tgc tgc gga ggc gcg gtt ttc ata ggt 48
 33 Met Asp Arg Glu Met Ala Ala Ser Cys Gly Gly Ala Val Phe Ile Gly
 34 1 5 10 15
 36 ctt gca ctc ttg acc ttg tca cca tac tat aaa gtg ctc ctc gct agg 96
 37 Leu Ala Leu Leu Thr Leu Ser Pro Tyr Tyr Lys Val Leu Leu Ala Arg
 38 20 25 30
 40 ctc ata tgg tgg tta cag tat tta atc acc aga gtc gag gcg cac ttg 144
 41 Leu Ile Trp Trp Leu Gln Tyr Leu Ile Thr Arg Val Glu Ala His Leu
 42 35 40 45
 44 caa gtg tgg atc ccc cct ctc aat gtt cgg gga ggc cgc gat gcc atc 192
 45 Gln Val Trp Ile Pro Pro Leu Asn Val Arg Gly Gly Arg Asp Ala Ile
 46 50 55 60
 48 atc ctc ctc acg tgc gca gtc cac cca gag cta atc ttt gac atc acc 240
 49 Ile Leu Leu Thr Cys Ala Val His Pro Glu Leu Ile Phe Asp Ile Thr
 50 65 70 75 80
 52 aaa ctc ctg ctc gcc ata ttc ggt ccg ctc atg gtg ctc cag gca ggc 288
 53 Lys Leu Leu Leu Ala Ile Phe Gly Pro Leu Met Val Leu Gln Ala Gly
 54 85 90 95

Does Not Comply
Corrected Diskette Reader

P2

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 11:19:37

Input Set : A:\13_082 Substitute Sequence listing US.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

56 ata acc aaa gtg ccg tac ttc gtg cgt gcg cag ggg ctc att cgt gcg 336
57 Ile Thr Lys Val Pro Tyr Phe Val Arg Ala Gln Gly Leu Ile Arg Ala
58 100 105 110
60 tgt atg ttg gtg cgg aag gct gcg ggg ggt cat tat gtc caa atg gcc 384
61 Cys Met Leu Val Arg Lys Ala Ala Gly Gly His Tyr Val Gln Met Ala
62 115 120 125
64 ttc atg aag cta gct gcg ctg aca ggt acg tac gtt tat gac cat ctc 432
65 Phe Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His Leu
66 130 135 140
68 act cca ttg cag gat tgg gcc cac gcg ggc cta cga gac ctt gca gtg 480
69 Thr Pro Leu Gln Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala Val
70 145 150 155 160
72 gcg gta gag ccc gtc atc ttc tct gac atg gag gtc aag atc atc acc 528
73 Ala Val Glu Pro Val Ile Phe Ser Asp Met Glu Val Lys Ile Ile Thr
74 165 170 175
76 tgg ggg gcg gac acc gcg gca tgc ggg gac atc att tca ggt ctg ccc 576
77 Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Ser Gly Leu Pro
78 180 185 190
80 gtc tcc gct cga agg gga agg gag ata ctc ctg gga ccg gcc gat aat 624
81 Val Ser Ala Arg Arg Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Asn
82 195 200 205
84 ttt gaa ggg cag ggg tgg cga ctc ctt gcg ccc atc acg gcc tac tcc 672
85 Phe Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ser
86 210 215 220
88 caa cag aca cgg ggc cta ctt ggt tgc atc atc acc agc ctc aca ggc 720
89 Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly
90 225 230 235 240
92 cgg gac aag aac cag gtc gag ggg gag gtt caa gtg gtc tcc acc gct 768
93 Arg Asp Lys Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr Ala
94 245 250 255
96 aca caa tct ttc ctg gcg acc tgc gtc aac ggc gtg tgt tgg act gtc 816
97 Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr Val
98 260 265 270
100 ttc cat ggc gcc ggc tca aag acc ttg gcc ggc ccc aaa ggc cca atc 864
101 Phe His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile
102 275 280 285
104 acc cag atg tac act aat gtg gac cag gac ctc gtc ggc tgg cag gcg 912
105 Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln Ala
106 290 295 300
108 ccc cct ggg gcg cgc tcc atg aca cca tgc acc tgc ggc agc tcg gac 960
109 Pro Pro Gly Ala Arg Ser Met Thr Pro Cys Thr Cys Gly Ser Ser Asp
110 305 310 315 320
113 ctc tat ttg gtc acg aga cat gcc gac gtc att ccg gtg cgc cgg cgg 1008
114 Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg
115 325 330 335
W--> 117
ggc gac agt agg ggg agc ctg ctc tcc ccc agg cct gtc tcc tac ttg 1056 Gly Asp Ser Arg Gly Ser
W--> 118 340 345 350
E--> 120 aag ggc tct tcg ggt ggc cca ctg ctc tgc cct tcg ggg cac gct gtg 1104
121 Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala Val

```

insert a hard return

RAW SEQUENCE LISTING

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 11:19:37

Input Set : A:\13_082 Substitute Sequence listing US.txt

Output Set: N:\CRF4\02122004\J017736C.raw

```

W--> 122          355          360          365
E--> 124 ggc atc ttc cgg gct gct gtg tgc acc cgg ggg gtt gea aaa gcg gtg 1152
      125 Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala Val
W--> 126          370          375          380
E--> 128 gac ttc ata cct gtt gag tct atg gaa act acc atg cgg act agt agc 1200
      129 Asp Phe Ile Pro Val Glu Ser Met Glu Thr Thr Met Arg Thr Ser Ser
W--> 130 385          390          395          400
E--> 132 gct tgg cgt cac ccg cag ttc ggt ggt taa 1230
      133 Ala Trp Arg His Pro Gln Phe Gly Gly *
E--> 134          405

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 02/12/2004
PATENT APPLICATION: US/10/017,736C TIME: 11:19:38

Input Set : A:\13_082 Substitute Sequence listing US.txt
Output Set: N:\CRF4\02122004\J017736C.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 117

VERIFICATION SUMMARY

DATE: 02/12/2004

PATENT APPLICATION: US/10/017,736C

TIME: 11:19:38

Input Set : A:\13_082 Substitute Sequence listing US.txt

Output Set: N:\CRF4\02122004\J017736C.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:117 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:17
L:118 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:120 M:254 E: No. of Bases conflict, LENGTH:Input:1104 Counted:1056 SEQ:1
L:122 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
M:254 Repeated in SeqNo=1
L:126 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:130 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:134 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:134 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1230 Found:1182 SEQ:1
L:1076 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1079 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1083 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:1084 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:1102 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:20
L:1103 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0